bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Multisite Phosphorylation and Binding Alter Conformational Dynamics of the 4E-BP2 Protein

Spencer Smyth[1,2,*], Zhenfu Zhang[1,2,*], Alaji Bah[3,+], Thomas E. Tsangaris[2], Jennifer Dawson[3,$], Julie D. Forman-Kay[3,4], and Claudiu C. Gradinaru[1,2,#]

[1]Department of Physics, University of Toronto, Toronto, Ontario, M5S 1A7, Canada; [2]Department of Chemical & Physical Sciences, University of Toronto Mississauga, Mississauga, Ontario, L5L 1C6, Canada; [3]Program in Molecular Medicine, The Hospital for Sick Children, Toronto, ON M5G 0A4, Canada; [4] Department of Biochemistry, University of Toronto, Toronto, ON M5S 1A8, Canada

[*] These authors contributed equally to this work

[+] Present address: Department of Biochemistry and Molecular Biology, SUNY Upstate Medical University, Syracuse, NY 13210, United States

[$] Present address: Genomic Medicine Institute, Cleveland Clinic, Cleveland, OH 44195, United States

[#] Author to whom correspondence may be addressed: claudiu.gradinaru@utoronto.ca

KEYWORDS: IDP, single-molecule fluorescence, anisotropy decay, PET-FCS, smFRET

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

ABSTRACT

Intrinsically disordered proteins (IDPs) play critical roles in regulatory protein interactions, but detailed structural/dynamics characterization of their ensembles remain challenging, both in isolation and they form dynamic 'fuzzy' complexes. Such is the case for mRNA cap-dependent translation initiation, which is regulated by the interaction of the predominantly folded eukaryotic initiation factor 4E (eIF4E) with the intrinsically disordered eIF4E binding proteins (4E-BPs) in a phosphorylation-dependent manner. Single-molecule Förster resonance energy transfer showed that the conformational changes of 4E-BP2 induced by binding to eIF4E are non-uniform along the sequence; while a central region containing both motifs that bind to eIF4E expands and becomes stiffer, the C-terminal region is less affected. Fluorescence anisotropy decay revealed a nonuniform segmental flexibility around six different labelling sites along the chain. Dynamic quenching of these fluorescent probes by intrinsic aromatic residues measured via fluorescence correlation spectroscopy report on transient intra- and inter-molecular contacts on ns-μs timescales. Upon hyperphosphorylation, which induces folding of ~40 residues in 4E-BP2, the quenching rates decreased at labelling sites closest to the phosphorylation sites and within the folded domain, and increased at the other sites. The chain dynamics around sites in the C-terminal region far away from the two binding motifs were significantly reduced upon binding to eIF4E, suggesting that this region is also involved in the highly dynamic 4E-BP2:eIF4E complex. Our time-resolved fluorescence data paint a sequence-level rigidity map of three states of 4E-BP2 differing in phosphorylation or binding status and distinguish regions that form contacts with eIF4E. This study adds complementary structural and dynamics information to recent studies of 4E-BP2, and it constitutes an important step towards a mechanistic understanding of this important IDP via integrative modelling.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

INTRODUCTION

Intrinsically disordered proteins (IDPs) are a class of proteins that lack well-defined three-dimensional structures while still carrying out biological functions (1-3). IDPs play a crucial role in mediating interactions with multiple partners and often function as protein interaction hubs (4, 5). IDPs within these protein complexes can undergo disorder-to-order transitions or remain dynamic (6). The lack of stable folded structures observed in IDPs leads to the simplistic assumption that IDPs resemble random coils. In fact, IDPs have transient secondary and tertiary structures and preferential backbone torsion angle propensities due to electrostatic and other interactions based on their amino acid sequence, and exhibit a wide range of compactness (6-9).

Cap-dependent initiation of translation is regulated by the interaction of eukaryotic initiation factor 4E (eIF4E) with disordered eIF4E binding proteins (4E-BPs) in a phosphorylation-dependent manner (10-12). The eIF4E protein, which binds the 7-methyl guanosine cap structure of mRNA at the 5' end, has been shown to be an oncogene and be involved in the induction of cellular transformation (13, 14). The eIF4G, a scaffolding protein, plays a crucial role in docking and assembling several components of the translation initiation machinery at the 5' cap of mRNA to recruit the ribosome (15). The 4E-BP2 protein is involved in controlling cell growth and proliferation via regulating mRNA translation (16) and in immunity to viral infections (17). Neural 4E-BP2 also functions in regulating synaptic plasticity, playing an essential role in learning and memory, and has been implicated in autism spectrum disorder (18, 19).

The interaction between eIF4E and eIF4G is the fundamental step that initiates the translation process. This interaction involves the canonical binding helix-forming YxxxxLϕ motif of eIF4G, which is also found in all 4E-BPs, binding to the same convex interface of eIF4E. Thus, the 4E-BPs regulate translation by competing with eIF4G to prevent the assembly of the eIF4F

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

complex and the subsequent mRNA recruitment to the ribosome. Binding of IDPs often leads to ordering, and transient helical structure around the canonical YxxxxLϕ motif is stabilized upon eIF4E binding. However, the 4E-BP2:eIF4E complex has been shown by NMR to be highly dynamic with an exchanging bipartite interface (20), in which the secondary binding site ^{78}IPGTV^{82} interacts with the lateral surface of eIF4E, as revealed by an X-ray crystal structure capturing a snapshot of the complex (21).

The 4E-BP2 protein is hierarchically phosphorylated. Modification of the first two sites T37 and T46 results in the hypo-phosphorylated state(22); this 2-site phosphorylation (2P) induces formation of a 4-stranded β-sheet structure from residues 18-62, partially sequestering the canonical binding motif and weakening eIF4E binding but still enabling competition with eIF4G(12). Additional phosphorylation at S65, T70, and S83 leading to a 5-site (5P) state stabilizes the fold (23), further decreasing the eIF4E affinity and allowing translation initiation to proceed (10). The disordered region C-terminal to the folded domain (C-IDR) remains disordered after phosphorylation and stabilizes the folded domain via long-range interactions (12, 23). However, important details of the structure and dynamics of 4E-BP2 and the eIF4E:4E-BP2 complex remain unknown, which prevent a clear mechanistic picture of the function of 4E-BP2 and its phosphoregulation of translation initiation.

Typically, IDPs have a wide range of interchanging conformations, therefore obtaining dynamic information is critically important for understanding their biological functions. Single-molecule fluorescence (SMF) techniques have been applied previously to measure the conformational heterogeneity, the global dimensions, and the dynamics of IDPs (24, 25). Here, we used fluorescence anisotropy decay (FAD) and fluorescence correlation spectroscopy (FCS) to characterize global and local peptide chain motions in 4E-BP2 upon multisite phosphorylation and

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

upon binding to eIF4E. FCS and FAD are sensitive to chain motions on the nanosecond-to-microsecond time scale, which is highly relevant for protein folding and for IDP dynamics (26). FAD measurements informed on the local, segmental flexibility of the peptide chain at various sites of the protein, leading to a coarse rigidity map of 4E-BP2 in the NP, 5P and eIF4E-bound states. Using FCS, we resolved up to three different timescales of intra-molecular conformational dynamics in 4E-BP2 under non-phosphorylated (NP), 5-site-phosphorylated (5P) and denaturing conditions. In addition, we obtained kinetic information (amplitude and lifetime) of key inter-molecular contacts of the dynamic binding interface between NP 4E-BP2 and eIF4E.

Single-molecule Förster resonance energy transfer (smFRET) measurements of two different regions of the protein delineated changes in intramolecular distances and chain rigidity upon multisite phosphorylation, and upon binding to eIF4E. Surprisingly, smFRET showed an increase in the distance between residues 32 and 91 upon phosphorylation, despite folding-induced compaction of residues 18 to 62 (12). While in complex with eIF4E, a region of 4E-BP2 containing both canonical and secondary binding sites expands and stiffens considerably, while the C-terminal region expands only slightly but remains highly flexible. Our multifaceted SMF characterization of the 4E-BP2 conformational dynamics is an important step towards understanding the interplay between folding and release of binding to eIF4E, and it provides valuable information for calculating conformational ensembles of this multistate IDP via integrative modelling.

RESULTS AND DISCUSSION

Local chain dynamics measured by FAD

The structural flexibility of IDPs is essential for regulating their interactions with other proteins and their role in signaling processes (3). FAD measures the rotational dynamics of the

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

emission dipole of a fluorophore and is therefore a suitable reporter of the local (segmental) chain flexibility around the labelling site. Inferred parameters from FAD analysis relate to the spatial confinement and the friction experienced by the dye, the movement of a protein segment around the labelling site, and the hydrodynamic radius of the protein segment (27). Probing local conformational dynamics is particularly relevant for IDPs, for which motions of different segments may be uncorrelated and obscured by the averaged global dynamics.

Six single-cysteine 4E-BP2 constructs were used for site-specific fluorescence labelling: C0 (i.e., C0ins/C35S/C73S mutations), C14 (S14C/C35S/C73S), C35 (C73S), C73 (C35S), C91 (C35S/C73S/S91C), and C121 (C35S/C73S/C121ins) (**Fig. S1, Table S1**). While the fastest, sub-ns FAD lifetime (ρ_{dye}) typically describes the rotation of the dye-linker around the labelling site, in IDPs with transient structure the slowest FAD lifetime describes the rotational diffusion of a protein segment (ρ_{seg}), where individual segments can rotate independently(28, 29). Alternatively, for folded proteins the slowest FAD lifetime describes the rotational diffusion of the entire protein.(30)

Fig. 1A shows a family of FAD curves of NP 4E-BP2 with the fluorophore attached to each of the six different mutated cysteine sites. As dye-protein interactions are expected to be negligible for Atto 488 (31) (**Fig. S3**), the variations in the anisotropy decays indicate that the chain flexibility is site/region dependent. FAD fitting parameters using **Eq. 1** (see Methods) are listed in **Table 1**, for each labelling site and each sample condition. **Fig. 1C** shows a comparison between segmental lifetimes (ρ_{seg}) around each 4E-BP2 labelling site in different phosphorylation states, i.e., NP, 2P and 5P. These values range from 1.4 ns to 3.4 ns and vary considerably with the labelling site and with the phosphorylation state.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

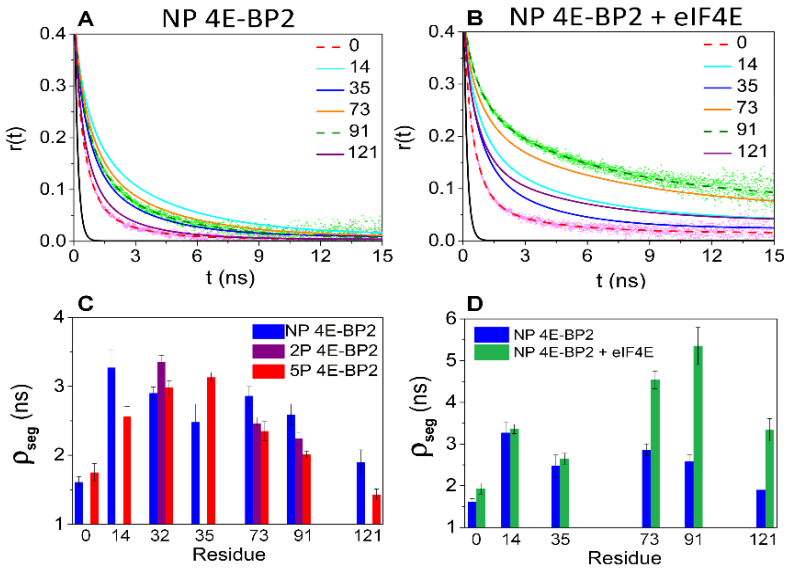

Figure 1. Fluorescence anisotropy decay (FAD) data for 4E-BP2 fit to **Eq. 1.** The 4E-BP2 protein was labelled with Atto 488 at positions 0, 14, 35, 73, 91 and 121 along the sequence, as described in the text. NP 4E-BP2 anisotropy data and fitted curves for all six labelling sites are shown in the absence (**A**) and in the presence of 1 µM of eIF4E (**B**); for reference the FAD of the free dye is shown in black in both panels. (**C**) The slowest rotational lifetime (ρ_{seg}) obtained by fitting FAD curves at each labelling site for NP (**blue**), 2P (**purple**) and 5P (**red**) 4E-BP2. (**D**) The slowest rotational lifetime (ρ_{seg}) obtained by fitting FAD curves at each labelling site for NP 4E-BP2 in the absence (**blue**) and in the presence of (**green**) of eIF4E.

Table 1. Anisotropy decay parameters for phosphorylated/bound 4E-BP2 states labelled at different sites[a].

	5P 4E-BP2		***2P 4E-BP2***		***NP 4E-BP2***		***NP 4E-BP2 + eIF4E***	
	ρ_{dye}	ρ_{seg} ***(a)***	ρ_{dye}	ρ_{seg} ***(a)***	ρ_{dye}	ρ_{seg} ***(a)***	ρ_{dye}	ρ_{seg} ***(a)***
C0	0.53 ns	1.75 ns *(0.72)*	-	-	0.51 ns	1.61 ns *(0.66)*	0.52 ns	1.93 ns *(0.70)*
S14C	0.56 ns	2.56 ns *(0.51)*	-	-	0.64 ns	3.27 ns *(0.42)*	0.62 ns	3.36 ns *(0.49)*
H32C	0.34 ns	2.98 ns *(0.59)*	0.41 ns	3.35 ns *(0.49)*	0.35 ns	2.90 ns *(0.55)*	-	-
C35	0.55 ns	3.13 ns *(0.41)*	-	-	0.64 ns	2.48 ns *(0.52)*	0.62 ns	2.65 ns *(0.51)*
C73	0.55 ns	2.35 ns *(0.54)*	0.39 ns	2.46 ns *(0.60)*	0.62 ns	2.86 ns *(0.47)*	0.61 ns	4.54 ns *(0.46)*
S91C	0.57 ns	2.01 ns *(0.51)*	0.33 ns	2.24 ns *(0.66)*	0.61 ns	2.59 ns *(0.50)*	0.65 ns	5.35 ns *(0.43)*
C121	0.52 ns	1.43 ns *(0.61)*	-	-	0.53 ns	1.90 ns *(0.60)*	0.59 ns	3.34 ns *(0.59)*

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

[a] All data were fit to **Eq. 1** to estimate two rotational correlation lifetimes (dye & segment) and their fractions (*a* is the fraction of the slowest component). Fitting error margins are on the order of ±0.1 ns for the lifetimes and ±0.05 for amplitudes **(Table S2).**

Many IDPs are more compact than an ideal statistical coil of the same length due to transient intra-molecular contacts (7, 8, 32). Slower segmental dynamics in PBS buffer (pH 7.4, 140 mM NaCl) was observed than in chemical denaturant (6M GdmCl) at all labelling sites (**Table S2**). The chemically denatured NP and 5P states of 4E-BP2 have similar segmental flexibility signatures, with $\rho_{seg} = 0.8 - 0.9$ ns at the ends and $\rho_{seg} = 1.3 - 1.6$ ns at internal sites. The values and the trend here match previous measurements for denatured proteins and the expectations for a random coil state(27, 29). Additionally, the pattern of FAD lifetimes is consistent overall with that of ^{15}N relaxation rates(20) and of ^{1}H-^{15}N nuclear Overhauser effect (NOE) values (12). This suggests that FAD-measured segmental dynamics can probe transient secondary structure and non-random chain contacts, and reports on the local degree of disorder in IDPs.

Another contribution to the heterogeneous segmental flexibility is the amino acid composition around each labelling site, with glycine and serine being the most flexible, and proline, isoleucine and valine the most rigid (33). With 4 Ser and 2 Gly among the first 10 residues, the N-terminus is the most flexible region of NP 4E-BP2. Considering a 10-residue window, the positions 14, 73 and 91 are flanked by several rigid residues (Pro, Ile, Val) and the protein appears to be much less flexible in these regions. The slowest rotational lifetime was observed at position 14 ($\rho_{seg} = 3.27$ ns), which, in addition to two Pro residues, has two positively charged Arg residues in its proximity, which may rigidify the segment further via electrostatic interactions.

In previous NMR studies, we have shown that a region of 4E-BP2 (residues 18 to 62) folds upon phosphorylation, while the rest of the chain remains disordered (12). We used FAD to probe

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

the changes in segmental flexibility of 4E-BP2 that are induced by partial (2P) and full (5P) phosphorylation (**Fig. 1C, Table 1**). The rotational lifetime (ρ_{seg}) increases at position 35 (and only slightly at 32, located within a long loop), while it decreases at all the other dye positions, which are outside the folded domain, indicating that the chain becomes more flexible. This is consistent with the formation of the four-stranded beta-sheet fold between residues 18-62 and with the C-terminal region remaining disordered after phosphorylation. From an entropic perspective, increased conformational flexibility near the secondary binding site (residues 78 to 82) may also contribute to decreasing the affinity for eIF4E.

FAD has been used previously to quantify and differentiate local binding constants of IDPs in the context of multisite interactions (28) . **Fig. 1B** shows anisotropy decay data and fitted curves for the 4E-BP2:eIF4E complex at each of the six labeling sites on 4E-BP2. In contrast with the apo sample, these curves decay to significantly higher asymptotic values (r_{inf}, **Table S2**), indicating that the local motions around each labeling site in 4E-BP2 are hindered after binding to eIF4E. The largest changes in chain flexibility occur at three C-terminal sites while the changes observed in the N-terminal sites are minor (**Fig. 1D**). At positions 73, 91 and 121, the segmental rotational lifetime ρ_{seg} nearly doubles, from 2-3 ns in the apo state to 4-5 ns in the bound state. Similarly, rotational freedom of the probe as measured by the half-cone angle, θ (**Eq. 2**), is significantly reduced at the C-terminal labelling sites (**Fig. S2**).

Positions 73 and 91 are located near the secondary binding site while position 121 is far from either binding site but the changes in lifetime are consistent with binding-induced changes to NMR intensity ratios(20), which demonstrate complete broadening for residues 45-88, significant broadening for residues 34-90, and broadening at residue 120. The data for positions 73 and 91 likely reflect favorable interactions at the secondary binding site, as well as potentially the

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

competitive interaction of the secondary binding region with the disordered N-terminus of eIF4E that has been suggested to act as a negative regulator.(34) The NMR broadening results for position 120 together with these FAD data on position 121 support a picture of the 4E-BP2:eIF4E dynamic complex involving a more extensive part of the C-terminus.

The results demonstrate that segmental motion parameters measured by FAD can be used to disentangle the binding contributions of different regions in an intrinsically disordered protein as it interacts with its binding partners. More specifically, the changes in θ are anti-correlated with the changes in ρ_{seg}, indicating that the rotational freedom of the probe and the segmental dynamics of 4E-BP2 are hindered in the bound state.

Non-local chain dynamics in 4E-BP2 measured by PET-FCS

The same Atto488-labelled 4E-BP2 constructs used for FAD (**Fig. S1, Table S1**) were also used for FCS experiments. FCS is sensitive to intensity fluctuations caused by the diffusion of the labelled protein and by the photophysics of the fluorophore (**Eq. 3)**. Notably, the fluorophore can be dynamically quenched by aromatic amino acids via photoinduced electron transfer (PET)(35, 36). Tryptophan (Trp) and tyrosine (Tyr) are the strongest PET quenchers, with a quenching range of 5-10 Å for typical fluorophores (37, 38). The two tyrosines in the 4E-BP2 sequence (Y34 and Y54) are most likely acting as quenchers in our FCS measurements.

Fig. 2A-B shows experimental FCS data, fitting curves, and residuals for a representative sample, 4E-BP2 labelled with Atto488 at position 73, in the NP state. Models with one diffusion component and two or three (faster) kinetic components (**Eq. 3**) satisfactorily fit the experimental autocorrelation decays measured for all 4E-BP2 samples. The (sub-diffusion) kinetic components in FCS data are attributed to intrinsic triplet-state kinetics of the probe and to dynamic PET quenching induced by the protein environment.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The FCS lifetimes obtained by fitting the data to **Eq. 3** can be grouped into three timescales: $\tau_1 \approx$ 50-500 ns, $\tau_2 \approx$ 3-30 µs, and $\tau_3 \approx$ 100 µs, respectively (**Fig. 2, Table 2**). Conversely, the free dye exhibits a single kinetic component with a lifetime of 12.9 ± 1.1 µs (**Fig. S2**), which is similar to a previously reported triplet lifetime of Atto488.(39) Thus, for 4E-BP2, the τ_1 and τ_3 lifetimes are attributed to PET quenching of the dye due to conformational dynamics of the protein, while the τ_2 lifetime may be a combination of intrinsic (triplet) and protein-induced (PET) photophysics. Chain motions on the sub-µs (τ_1) timescale are typically associated with interconversion of states within the disordered conformational ensemble and with proximal loop formation.(40, 41) The intermediate (τ_2) timescale is similar to that of forming transient secondary structure and distal chain contacts (42), while the slowest (τ_3) kinetics is on the same timescale as concerted motions of the protein chain, such as domain movements or transient tertiary structural contacts (43, 44).

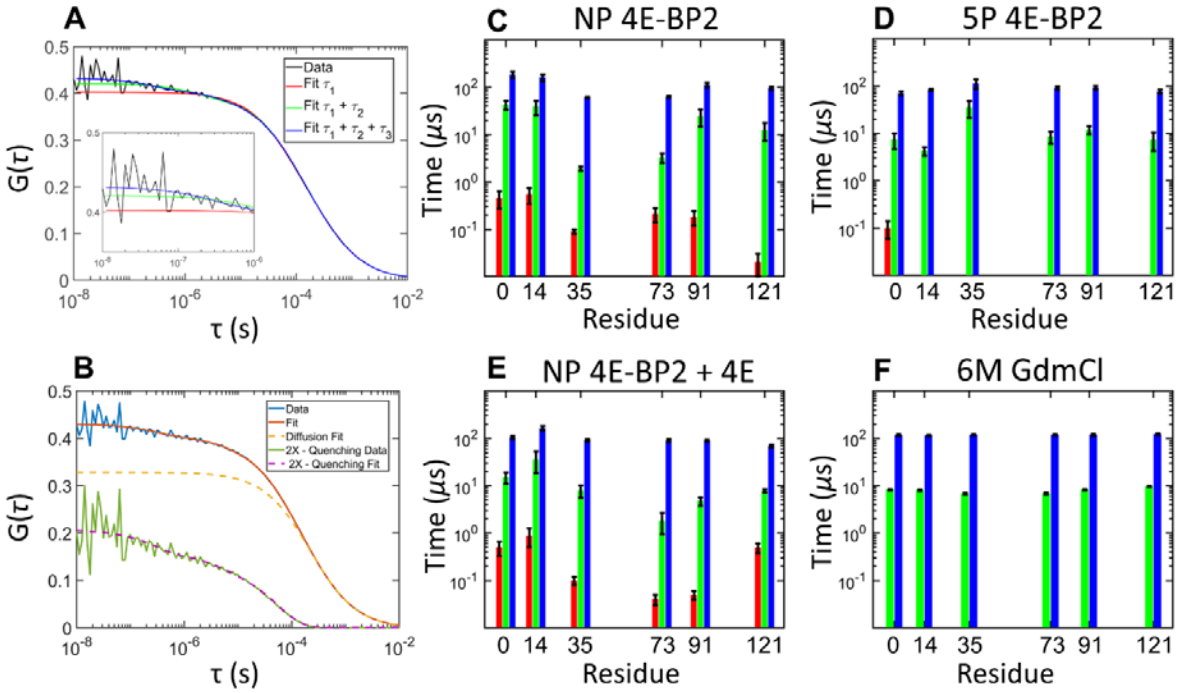

Figure 2. FCS lifetimes of 4E-BP2 labelled with Atto488 at residue 0, 14, 35, 73, 91 or 121. Experimental curves were fit to **Eq. 3**. NP 4E-BP2 labeled at residue 73 was fitted with 1, 2 or 3 kinetic decay components (**A**), with the kinetic decays and diffusion for the best fit plotted separately (**B**). FCS experiments were performed on 4E-BP2 in different states: non-phosphorylated (**C**), hyper-phosphorylated (**D**), in the presence of 0.5 µM of elF4E (**E**), and in the presence of 6 M GdmCl (**F**). Different decay timescales are

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

shown in different colors: τ_1=0.03-1 µs (**red**), τ_2=3-30 µs (**green**) and τ_3≈100 µs (**blue**). The fitting error bars for each lifetime are shown in the figure. The full list of fitting parameters is given in **Table 2**.

For NP 4E-BP2, three kinetic components were resolved at each labelling position, which follow the $\boldsymbol{\tau_1}$, $\boldsymbol{\tau_2}$ and $\boldsymbol{\tau_3}$ timescales. The lifetimes and amplitudes are highly site-dependent, especially for the two fastest components (**Fig. 2C**). These results show higher flexibility in the central region of the protein that folds upon phosphorylation, and in the C-terminal region, in agreement with previous NMR data(12, 23). Similar PET-FCS analysis, with multiple lifetimes of intrachain dynamics between ~100 ns and ~100 µs, was previously reported for other IDPs, e.g., the N-terminal domain of p53-TAD (42) and the mouse prion protein moPrP (43).

Table 2. FCS decay parameters for five-phospho, non-phospho and elF4E-bound 4E-BP2 [a].

	5P-BP2			**NP-BP2**			**NP-BP2 + elF4E**		
	τ_1(µs)	**τ_2(µs)**	**τ_3(µs)**	**τ_1(µs)**	**τ_2(µs)**	**τ_3(µs)**	**τ_1(µs)**	**τ_2(µs)**	**τ_3(µs)**
C0	0.10 *(0.17)*	7.4 *(0.05)*	70.2 *(0.02)*	0.45 *(0.06)*	42.9 *(0.20)*	183.9 *(0.21)*	0.50 *(0.07)*	15.2 *(0.10)*	104.9 *(0.24)*
S14C	-	4.3 *(0.06)*	83.9 *(0.24)*	0.54 *(0.07)*	38.8 *(0.16)*	161.6 *(0.22)*	0.88 *(0.08)*	35.5 *(0.11)*	164.4 *(0.26)*
C35	-	34.8 *(0.08)*	110.9 *(0.18)*	0.09 *(0.10)*	1.9 *(0.08)*	61.4 *(0.16)*	0.10 *(0.15)*	7.9 *(0.06)*	91.8 *(0.25)*
C73	-	8.4 *(0.05)*	90.9 *(0.22)*	0.21 *(0.07)*	3.3 *(0.05)*	63.3 *(0.16)*	0.04 *(0.36)*	1.8 *(0.04)*	89.0 *(0.24)*
S91C	-	11.9 *(0.08)*	92.5 *(0.19)*	0.18 *(0.08)*	24.7 *(0.09)*	112.0 *(0.25)*	0.05 *(0.28)*	4.8 *(0.08)*	90.0 *(0.24)*
C121	-	7.3 *(0.05)*	76.1 *(0.19)*	0.02 *(0.24)*	12.5 *(0.07)*	95.7 *(0.25)*	0.49 *(0.22)*	7.8 *(0.58)*	68.4 *(0.29)*

[a] All data were fit to **Eq. 3** Lifetimes are given in microseconds, with corresponding amplitudes between brackets. Fitting errors are given in SI (Table S3).

In contrast to the NP state, the 5P state exhibits only two FCS kinetic components ($\boldsymbol{\tau_2}$ and $\boldsymbol{\tau_3}$) at each labelling position, except for the C0 which required three (**Fig. 2B**). Likewise, the fast

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(τ_1) kinetic component was not resolved in the chemically denatured state of 4E-BP2 (**Fig. 2D**). Surprisingly, the 5P state and the denatured state show similar lifetimes along the sequence, with the denatured state having a more uniform distribution than 5P. For both samples, the τ_1 kinetics may have accelerated to timescales faster than 50 ns, such that it is not resolvable in our measurements. This is likely the case for the denatured state, as for protein L the rate of intramolecular contact is ~20 times faster in 6 M GdmCl(45). However, for the 5P state, the addition of phosphate groups and the induced folding in the 18-62 region probably slows down the protein reconfiguration time sufficiently to be "absorbed" into the τ_2 kinetic component, as previously reported for p53-TAD (42).

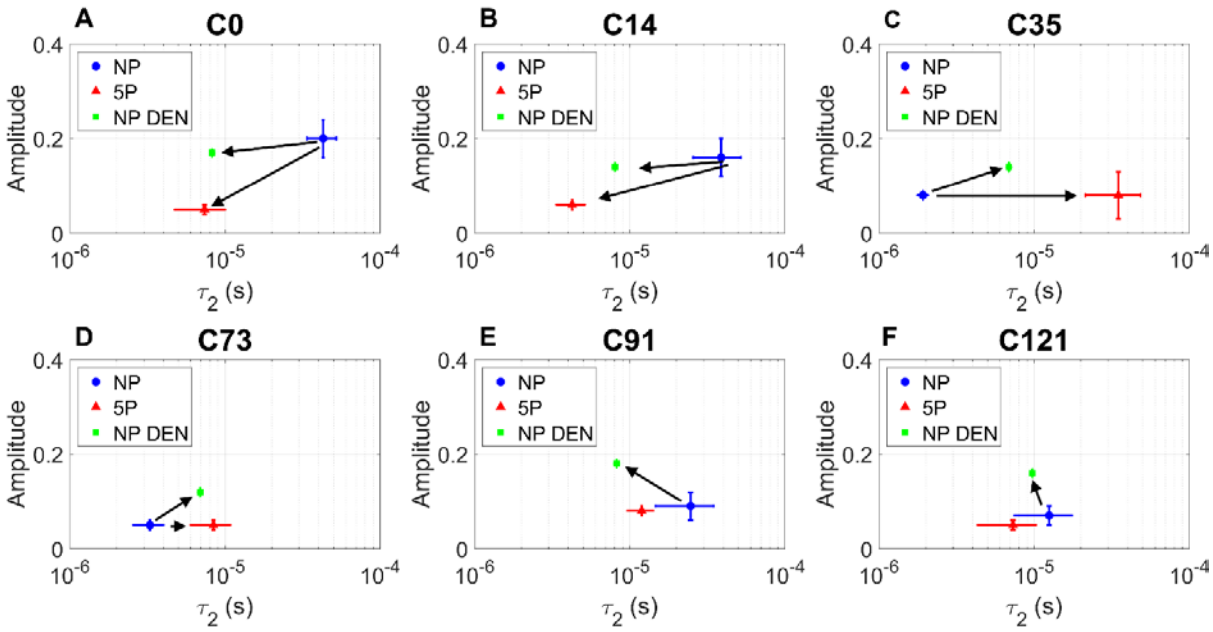

Figure 3. The changes in lifetime (τ_2) and amplitude (a_2) of the intermediate FCS kinetic component induced by multisite phosphorylation or by chemical denaturation (6 M GdmCl). 4E-BP2 was labelled with Atto488 at positions 0, 14, 35, 73, 91 and 121 along the sequence, and the experimental FCS curves were fit to **Eq. 3**. Different protein states are indicated as follows: non-phospho (**NP, blue circle**), five-phospho (**5P, red triangle**), and non-phospho denatured (**NP DEN, green square**). The fitting error bars for each parameter are indicated in the figure; all fitting parameters are given in **Table 2**.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Upon phosphorylation, the τ_2 kinetics speeds up at positions 0, 14 and 91 and it slows down at positions 35 and 73; the associated kinetic amplitude a_2 decreases at sites 0 and 14 and does not change at the other sites (**Fig. 3, Table 2**). Sites 35 and 73 are most proximal to the sites of phosphorylation (i.e., T37, T46, T70, S65 and S83), from which it can be inferred that the dominant effect of phosphorylation is to slow down intrachain contacts causing PET quenching. Similarly, slower conformational dynamics upon multisite phosphorylation has been observed for the disordered p53-TAD protein, even though phosphorylation does not lead to folding in this system.(42) In 4E-BP2, phosphorylation stabilizes the β-strand structure between residues 18 and 62, which is expected to limit the access of the fluorophore to tyrosine quenchers.(12)

For the denatured state of NP 4E-BP2 (in 6 M GdmCl), only the τ_2 and τ_3 kinetic components were observed (**Fig. 2D**); nearly identical results were obtained for denatured 5P 4E-BP2 under the same GdmCl conditions (**Table S3**). The fitted lifetimes and their respective amplitudes showed little variance along the chain, e.g., τ_2 = 7-10 μs and τ_3 = 100-120 μs. This highlights the influence that residual transient secondary and tertiary structure have on the conformational ensemble of disordered states and on the intrachain dynamics of 4E-BP2. The changes observed are consistent with denaturants (good solvents) reducing the strength of intramolecular interactions (hydrophobic, hydrogen bonds, van der Waals) in the context of protein-protein and protein-solvent interactions.[5,9]

Given the tight binding affinity of 4E-BP2 for eIF4E (k_d = 3.2 ± 0.6 nM)(12), saturating amounts of eIF4E (0.5 μM) were used to ensure that nearly all (>99%) 4E-BP2 molecules were in the bound state. FCS measurements on Atto488-labelled 4E-BP2 w/wo unlabeled eIF4E (**Fig. S3**) showed a shift to longer diffusion times (larger R_H) in the presence of eIF4E, which is consistent with the formation of the 4E-BP2:eIF4E complex. The R_H values estimated by fitting the data to

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Eq. 3 are 29.0 ± 0.5 Å for the NP 4E-BP2 (independent of labelling site) and 36.6 ± 0.8 Å for the 4E-BP2:eIF4E complex. As reference, FCS measurements of Atto488 (5 nM) in the presence of excess eIF4E (1 μM) showed that the dye does not interact/bind with eIF4E (**Fig S3**).

When bound to eIF4E, NP 4E-BP2 exhibits three FCS kinetic components which fall into the same τ_1, τ_2, and τ_3 range as in the free (Apo) state. eIF4E (PDB: 3AM7) contains 8 Trp and 6 Tyr residues, of which four Trp (W46, W56, W73 and W102) and three Tyr (Y34, Y76, and Y 145) residues are largely surface exposed (**Table S4**). Therefore, the observed changes in PET quenching dynamics are the combined result of the new inter-molecular interactions between 4E-BP2 and the surface of eIF4E and the changes in intra-molecular contacts within 4E-BP2.

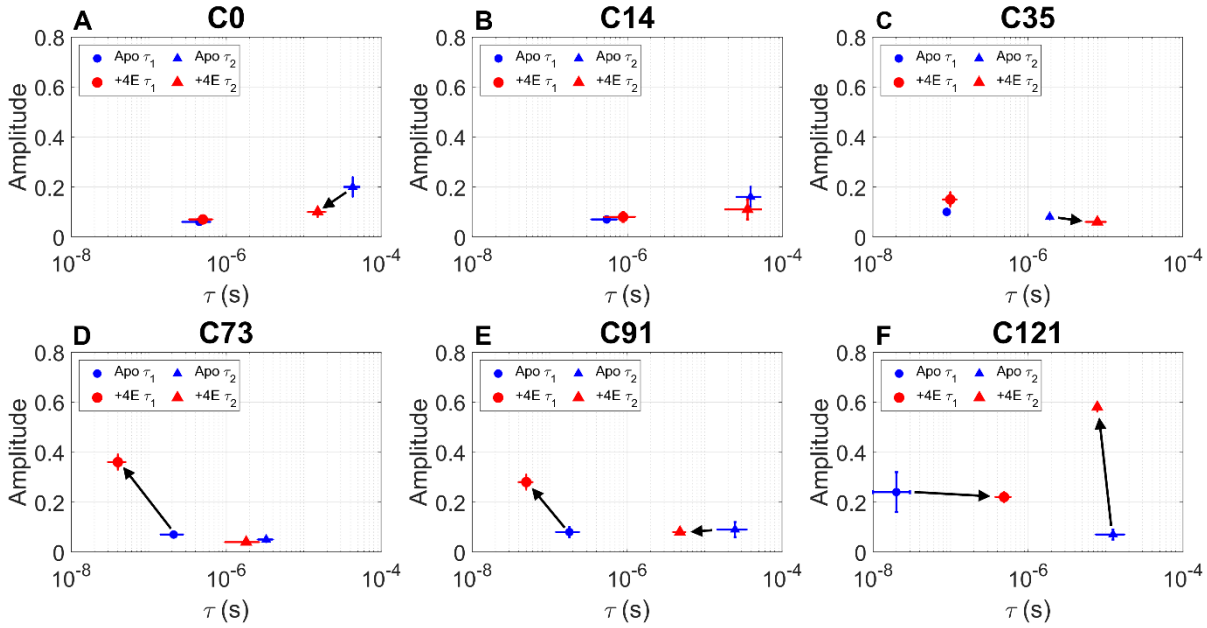

Figure 4. The changes in lifetime and amplitude of the two fastest FCS kinetic components (**τ_1 – circle**, and **τ_2 – triangle**) induced by binding of NP 4E-BP2 to eIF4E. FCS measurements on NP 4E-BP2 labelled with Atto488 at six different sites were performed in the Apo state (**blue**) and in the presence of 0.5 μM eIF4E (**red**), and the data were fit to **Eq. 3.** The fitting error bars for each parameter are indicated in the figure; all fitting parameters are given in **Table 2**.

4E-BP2 interacts with eIF4E through dynamic interactions involving at least an α-helical structure at the canonical ^{54}YXXXXLΦ^{60} motif and the secondary ^{78}IPGVT^{82} site (20, 21). Positions 73 and 91 show a decrease in lifetime accompanied by an increase in amplitude of the

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

fast component, and a slight decrease in lifetime of the intermediate component upon binding to eIF4E (**Fig. 4 D-E**). As such, τ_1 decreases from ~200 ns to 40-50 ns upon binding, while the decrease in τ_2 is more prominent at position 91, i.e., from ~25 μs to ~5 μs (**Table 2**). The changes at the C-terminus are also significant, with increases of the τ_1 lifetime and of the amplitude of the τ_2 component (**Fig. 4 F**). The decrease in lifetime and increase in amplitude of the τ_2 components of positions 73 and 91 are likely caused by dynamic exchange of the secondary binding site leading to quenching by aromatic residues on the surface of eIF4E. The involvement of residues 73, 91 and 121 in eIF4E binding is consistent with binding-induced NMR intensity changes reported previously(20). The changes at residue 121 upon binding mirror those observed by FAD (**Fig. 1 D**), providing further support for a dynamic complex involving more of the C-terminal portion of 4E-BP2. Taken together, the results highlight the dynamic nature and range of timescales present in the bound state and ability of PET-FCS to probe intermolecular dynamics within the bound state.

Chain dimensions and stiffness of 4E-BP2 assessed by smFRET

smFRET is exquisitely suited to delineate heterogeneous and dynamic states that are inherent to IDPs. smFRET can resolve heterogeneous population distributions and kinetics, is compatible with a wide range of solution conditions, and overcomes the ensemble averaging of established structural techniques such as NMR and small-angle X-ray scattering (SAXS)(25). By recording the arrival time, the color and the polarization of each detected photon, multiparameter fluorescence (MPF) detection permits access to additional intrinsic properties of fluorescence that can be related to properties of the conjugated molecule (46). The chain dimensions and stiffness of 4E-BP2 were assessed using smFRET with MPF on two double-cysteine FRET constructs, which were labelled stochastically with Alexa Fluor 488 (donor) and Alexa Fluor 647 (acceptor).

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

smFRET histograms for H32C/S91C in the NP, 5P, and NP+eIF4E states are shown in **Fig. 5**. Although each histogram was satisfactorily fit to a single Gaussian, a single FRET peak does not necessarily reflect a homogenous distribution of states. Indeed, the underlying population could be in fast exchange compared to the burst duration (~1 ms). The data suggests that the NP state is more compact than a statistical coil ensemble, i.e., $\langle E_{32-91}^{NP} \rangle_{exp} = 0.63 \pm 0.02$ vs. $\langle E_{32-91}^{NP} \rangle_{coil} = 0.43 \pm 0.03$ (see SI section 2.3). This is consistent with the presence of transient secondary structure observed by NMR(20), and with the multiple transient intrachain contacts observed by PET-FCS.

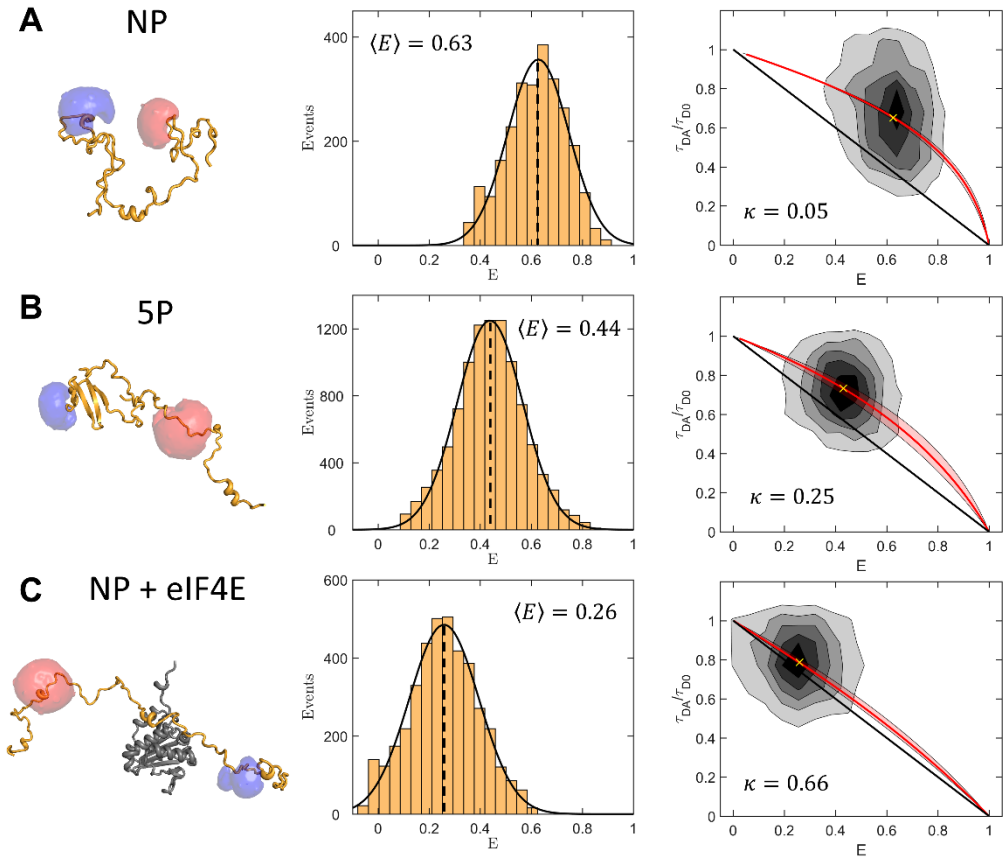

Figure 5. smFRET MPF results for the H32C/S91C 4E-BP2 construct labelled with Alexa Fluor 488 and Alexa Fluor 647 in the (**A**) non-phospho (NP), (**B**) five-phospho and (5P) (**C**) non-phospho eIF4E-bound (NP+eIF4E) states. (1st column) Cartoon depictions of conformations of 4E-BP2 in different states with representations of fluorophore accessible volume simulations (47). (2nd column) FRET efficiency histograms fit to a Gaussian function, with the dashed lines indicating the mean values. (3rd column) 2D histogram plots of the FRET efficiency (*E*) and the donor-only normalized fluorescence lifetime; τ_{DA} and

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

τ_{D0} are donor lifetimes in the presence and absence of acceptor, respectively. The black line shows the relation expected for a static rigid molecule. The red line is a relation for a worm-like chain with a stiffness parameter (κ) that passes through the centroid of a 2D gaussian fit indicated by a yellow cross. The red shaded region shows the uncertainty range of κ.

Upon multi-site phosphorylation 4E-BP2 undergoes folding to a beta-sheet domain between residues 18 and 62 (12). Typically, when proteins fold, FRET efficiency increases, following the overall compaction of the structure and contraction of most (but not all!) inter-residue distances(48, 49). In this case however, the FRET efficiency of the 5P state is lower for the dye pairs at residues 32 and 91, $\langle E_{32-91}^{5P} \rangle_{exp} = 0.44 \pm 0.02$, not higher than the NP state, $\langle E_{32-91}^{NP} \rangle_{exp} = 0.63 \pm 0.02$ (**Table 3**). The dye at C32 is positioned in the long loop between strands β1 and β2 which is proximal to the C-terminal end of the domain (PBD ID 2MX4), from where the C-IDR, containing the dye at C91, extends. An alternate construct designed to specifically probe the folded region (with both dyes within the folded domain boundaries from 18-62) would likely lead to an increase in FRET efficiency upon phosphorylation. The H32C/S91C construct, however, also partially probes the C-IDR, which expands when 4E-BP2 is phosphorylated (**Fig. 5B**).

The FRET efficiency of the NP H32C/S91C construct exhibits an even larger decrease upon binding to eIF4E, i.e., $E_{32-91}^{NP+4E} = 0.26 \pm 0.02$. This construct flanks both primary and secondary binding sites and should be sensitive to 4E-BP2:eIF4E interactions and the 3D arrangement of the complex. The large FRET decrease could result from a combination of more extended conformations being compatible with the eIF4E bound state, in which 4E-BP2 wraps around eIF4E (with the exception of the canonical helical element), and an excluded-volume effect exerted by eIF4E on 4E-BP2. The conformations of 4E-BP2 bound to eIF4E are thought to be defined by dynamic interactions of the canonical and secondary binding elements, and by significant structural disorder elsewhere (20).

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

smFRET histograms of the C73/C121 construct for the same three 4E-BP2 states are shown in **Fig. 6**. The average FRET efficiency for NP 4E-BP2, $\langle E^{NP}_{73-121}\rangle_{exp} = 0.58 \pm 0.02$, is also higher than that expected from a statistical coil, $\langle E^{NP}_{73-121}\rangle_{coil} = 0.49 \pm 0.03$, suggesting that the C-terminal disordered region contains transient intrachain contacts. The average C73/C121 FRET efficiency decreases to $\langle E^{5P}_{73-121}\rangle = 0.44 \pm 0.02$ upon phosphorylation, but only to $\langle E^{NP+4E}_{73-121}\rangle = 0.51 \pm 0.02$ upon binding to eIF4E. The electrostatic repulsion between the five phosphates may contribute to the C-IDR expansion of 5P 4E-BP2. This expansion of the C-IDR has also been observed from previously published NMR paramagnetic relaxation enhancement (PRE) data and chemical shift-derived measures of secondary structure.(23)

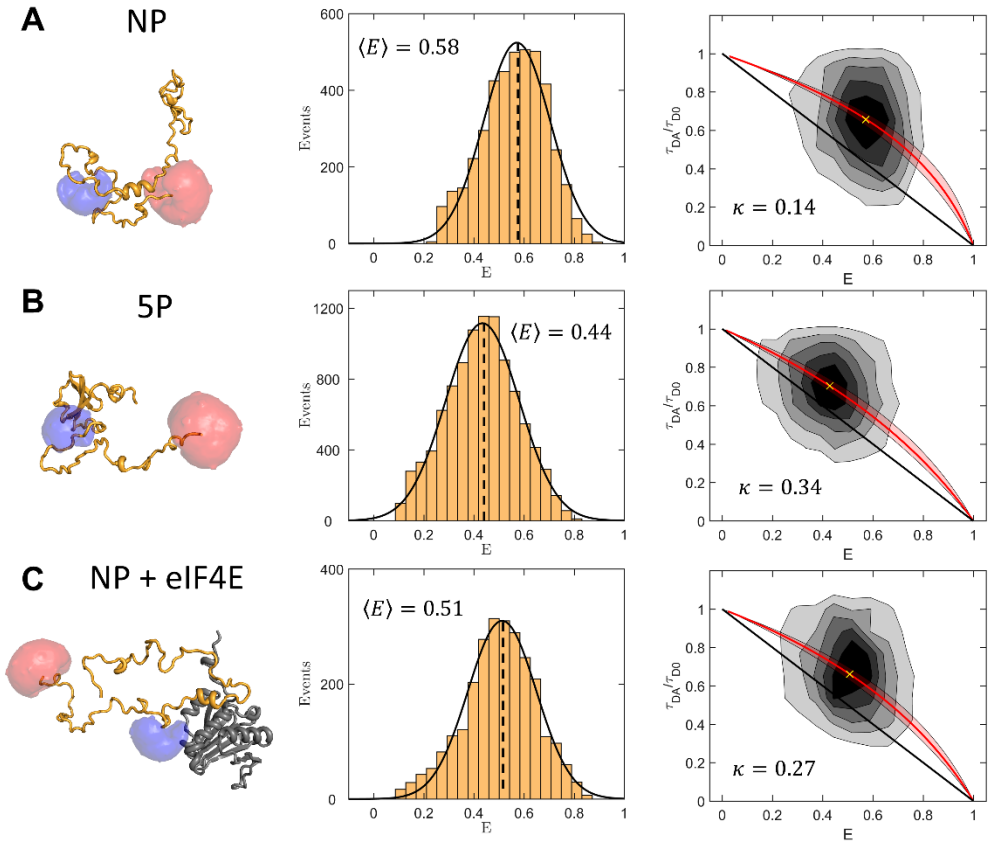

Figure 6. smFRET MPF results for the C73/C121 4E-BP2 construct labelled with Alexa Fluor 488 and Alexa Fluor 647 in the (**A**) non-phospho (NP), (**B**) five-phospho (5P) and (**C**) non-phospho eIF4E bound states (NP+4E). (1st column) Cartoon depictions of conformations of 4E-BP2 in different states with representations of fluorophore accessible volume simulations (47). (2nd column) FRET efficiency

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

histograms fit to a Gaussian function, with the dashed lines indicating the mean values. (3rd column) 2D histograms plots of the FRET efficiency (*E*) and the donor-only normalized fluorescence lifetime; τ_{DA} and τ_{D0} are donor lifetimes in the presence and absence of acceptor, respectively. The black line shows the relation expected for a static rigid molecule. The red line is a relation for a worm-like chain with a stiffness parameter (κ) that passes through the centroid of a 2D gaussian fit indicated by a yellow cross. The red shaded region shows the uncertainty range of κ.

For the bound state, the reduction in FRET is much less for C73/C121 than for H32C/S91C, i.e., $\Delta E_{73-121}^{NP\rightarrow+4E} = -(0.07 \pm 0.03)$ *vs.* $\Delta E_{32-91}^{NP\rightarrow+4E} = -(0.37 \pm 0.03)$. This suggests that C-terminal disordered conformations are less restricted in the 4E-BP2:eIF4E complex than those of the 32-91 region, as expected from the complete loss or very significant broadening of NMR resonances upon binding from residues 34-90 (20).

The relation between the donor lifetime and the FRET efficiency can be used to infer information about the dynamic exchange of the underlying states (50). A linear relation is expected for a static structure, with conformations that are rigid or fluctuate on a timescale slower than ~100 μs; in contrast, a nonlinear relation is expected for IDPs, as inter-photon times (1-10 μs) are much larger than typical chain reconfiguration times (~100 ns) (51). A family of dynamic τ *vs.* *E* lines based on a worm-like chain (WLC) model with variable stiffness κ (or persistence length) was generated using a method described by Barth et al. (52) The center of the experimental 2D FRET histogram was best matched to a WLC curve to infer the average stiffness for different regions of 4E-BP2 in different phospho/binding states (**Table 3**).

Table 3. Mean smFRET efficiency and WLC stiffness values of the non-phospho (NP), five-phospho (5P) and elF4E-bound 4E-BP2 (NP+4E).

	NP-BP2		**5P-BP2**		**NP-BP2 + elF4E**	
	<*E*>	**κ**	**<*E*>**	**κ**	**<*E*>**	**κ**
H32C/ S91C	0.63 ± 0.02	0.05 ± 0.03	0.44 ± 0.02	0.25 ± 0.10	0.26 ± 0.02	0.66 ± 0.17
C73/ C121	0.58 ± 0.02	0.14 ± 0.08	0.44 ± 0.02	0.34 ± 0.10	0.51 ± 0.02	0.27 ± 0.08

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

The stiffness parameter of the 32-91 region, $\kappa_{32-91}^{NP} = 0.05 \pm 0.03$, corresponds to a persistence length of $l_{p_{32-91}}^{NP} = 1.0 \pm 0.6$ nm, which is similar to $l_p = 0.4 \pm 0.07$ nm reported for a set of disordered and unfolded proteins.(53) The stiffness increases to $\kappa_{32-91}^{5P} = 0.25 \pm 0.10$ upon phosphorylation, and even more significantly, to $\kappa_{32-91}^{NP+4E} = 0.66 \pm 0.17$ upon binding to eIF4E (**Fig. 5** and **Table 3**). The larger stiffness in the 5P state is consistent with the appearance of a stable beta-fold between residues 18 and 62. In the bound state, the 2D FRET population lies close to the static line; this is consistent with the decreased dynamics expected as residues 49-67 form a predominantly stable helix when 4E-BP2 is bound to eIF4E.(20, 54) At the same time, the separation between the static and dynamic regimes is much reduced in the low *E* range, which increases the uncertainty in estimating the chain stiffness.

The 73-121 region has a stiffness of $\kappa_{73-121}^{NP} = 0.14 \pm 0.08$ and $\kappa_{73-121}^{5P} = 0.34 \pm 0.10$ in the NP and 5P states, respectively. While slightly larger, these values mirror the changes observed for the 32-91 region. The stiffening of the C-IDR upon phosphorylation can be attributed to stabilizing interactions with the folded domain established previously (23). However, in contrast to the 32-91 region, the stiffness of the C-terminal region shows only a moderate increase in the eIF4E-bound state, to $\kappa_{73-121}^{NP+4E} = 0.27 \pm 0.08$, as the 2D FRET histogram remains well separated from the static line (**Fig. 6**). Together with the FAD and PET-FCS results, this suggests that the 73-121 region remains disordered and interacts with eIF4E as part of a more extensive dynamic complex than the current bipartite model (20).

CONCLUSIONS

Dynamics is likely the key factor in understanding how 4E-BP2 regulates cap-dependent translation via interactions with the initiation factor eIF4E. Complementary to our previous NMR

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

studies of the 4E-BP2 protein (12, 20, 23), a suite of multiparameter fluorescence techniques was used here to define local and global conformational dynamics of this intriguing IDP in its various states. Polarization anisotropy (FAD) measurements showed evidence of heterogeneous chain flexibility within the 4E-BP2 sequence, which correlates with variations in transient secondary structure and local amino acid composition. Multisite phosphorylation decreased segmental flexibility only in the region that undergoes folding (residues 18-62), while it made the rest of the chain more flexible. As anticipated, in the presence of eIF4E, segmental dynamics near the two binding sites was slowed down by binding interactions. Surprisingly however, the C-terminal region adjacent to the secondary binding site experienced the largest changes, while the N-terminal region exhibited much less change. These results implicate a much larger region of 4E-BP2 interacting with eIF4E than previously thought, and showcase FAD as a sensitive method to probe local binding affinity, at the level of individual protein segments, in the context of multivalent interactions.

As As seen previously for other IDPs (42), the quenching rates decreased at labelling sites closest to the phosphorylation sites and within the folded domain, and increased at the other sites. This likely arises from a combination of local steric and electrostatic effects of the phosphate groups and the overall steric effect of the folded domain. In the presence of eIF4E, the quenching rates near the two binding motifs in 4E-BP2 increased significantly, as expected from increased contacts with several exposed Trp and Tyr residues along the extensive binding interface of eIF4E. Although contributions from intra- and inter-molecular contact could not be distinguished, the significant changes observed at the C-terminus support the involvement of the C-terminal region of 4E-BP2 in dynamic interactions with eIF4E.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

smFRET measurements informed on changes in 4E-BP2 chain dimensions and stiffness of central and C-terminal regions of the protein following phosphorylation or binding. In the NP state, as a consequence of transient intrachain secondary and tertiary contacts, the donor-acceptor distance turned out smaller (higher FRET efficiency) than the random coil prediction for both FRET constructs. In the 5P state, the separation between residues 32 and 91, which partially contains the folding domain, increased, instead of the decrease typically expected for a folding transition. When bound to eIF4E, the chain expands and stiffens considerably around the canonical and secondary binding sites, while the C-terminal region remains highly flexible. The canonical binding motif and secondary binding site provide specificity, while the dynamic nature of the complex which retains significant chain flexibility minimizes the entropic penalty. All phosphorylation sites except S65, which is phosphorylated last, are in regions that remain highly dynamic in the 4E-BP2:eIF4E complex. This facilitates the access of large kinases and allows a rapid response to cellular conditions. Our fluorescence-based characterization of the 4E-BP2 conformational dynamics is an important step towards understanding the interplay between folding and release of binding to eIF4E and its regulatory function, and it provides a foundation for future studies of IDP conformational and binding equilibria.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

MATERIALS AND METHODS

Materials. The fluorescent probes used for labelling the 4E-BP2 protein for SMF experiments were: Alexa 488 (A488) maleimide, Alexa 647 (A647) maleimide (ThermoFisher Scientific, Canada) and Atto 488 (At488) maleimide (ATTO-TEC GmbH, Germany). Guanidinium chloride (GdmCl) (G9284, Sigma Aldrich) was used for protein denaturation. All samples were diluted in phosphate-buffered saline (PBS) containing 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 1.8 mM KH_2PO_4 at pH 7.4. GdmCl solutions were adjusted to pH 7.4 for all the denaturation experiments.

Protein expression and purification. Small ubiquitin-like modifier (Sumo) fusion constructs of both the 4E-BP2 protein (residues 1 through 120) and eIF4E protein (residues 1 through 217) were expressed and purified as described previously(23). Briefly, the proteins were expressed in BL21-codonplus (DE3)-RIPL competent *E. coli* cells (Agilent Technologies) in Lysogeny broth at 37 ºC until OD_{600}~0.6-0.8, induced with isopropyl β-D-1-thiogalactopyranoside (IPTG), and expressed at 16 ºC for ~16 h. Protein was purified from cell lysate with a nickel–nitrilotriacetic acid (Ni–NTA) column followed by cleavage of the Sumo solubility tag with ULP1 at 4 ºC for ~16 h. The Sumo tag was separated using an N-NTA column followed by HiLoad Superdex 75 PG gel filtration column (28-9893-34, GE Healthcare) if the protein was not pure as assessed by SDS-PAGE. The molecular mass and the purity of protein samples were verified by electrospray ionization mass spectrometry (ESI-MS).

Phosphorylation of 4E-BP2 with activated Erk2 using a dialysis technique was performed as described previously(12, 23). Briefly, 50 mL of 5 µM Erk2 and 20 µM of 4E-BP2 were dialyzed in a 3 kDa MWCO dialysis bag in 1 L of buffer. The dialysis buffer contained 50 mM Tris-HCl, pH 7.5, 1 mM EGTA, 5 mM BME, 20 mM $MgCl_2$, and 10 mM EDTA, dialysis was performed at

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

20 ºC for 1-3 days. Phosphorylated 4E-BP2 was purified from Erk2 using a Ni-NTA column. Purity and degree of phosphorylation of 4E-BP2 was confirmed by ESI-MS.

All single cysteine proteins (C0/C35S/C73S (cysteine insertion at 0 position), S14C/C35S/C73S, C35/C73S, C73/C35S, C35S/C73S/S91C, C35S/C73S/C121 (insertion at 121 position)) were labeled by adding the At488 maleimide fluorophore to a 50 μL solution of 100 μM protein at a dye:protein molar ratio of 3:1. The double-cysteine mutants (H32C/C35S/C73S/S91C) and (C35S/insC121) was labeled with A488 maleimide and A647 maleimide by adding A488 and A647 to a 50 μL solution of 100 μM protein at a A488:A647:protein molar ratio of 1.3:3:1. TCEP was added at a 10× molar excess to the protein in order to reduce the disulfide bonds. All the maleimide-cysteine coupling reactions were performed in a PBS buffer at pH 7.4. Oxygen was removed by flushing the sample with argon gas in a desiccator for 5 min. The vial was capped tightly and shaken gently for 3 hours at room temperature. The excess dye was removed by size-exclusion chromatography using Sephadex G-50 gels (G5080, Sigma Aldrich) in a BioLogic LP system (731-8300, Bio-Rad).

All samples were diluted to concentrations of 1–10 nM and 20–50 pM, which are most suitable for FCS/FAD and smFRET burst experiments, respectively. For a typical experiment, a sample solution of 30 μL was dropped on the surface of plasma-cleaned coverslip. Non-specific protein adsorption to the coverslip was prevented by adding 0.005% (v/v) Tween-20 (P2287, Sigma-Aldrich) to the solution, and bovine serum albumin (BSA) (15260-037, ThermoFisher Scientific) was used to coat the clean coverslips. All experiments were performed at 20 °C.

Instrumentation. smFRET measurements were performed on a custom-built multiparameter fluorescence microscope (32). The donor was excited at 480 nm by frequency

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

doubling the infrared output of a femtosecond laser (Tsunami HP, Spectra Physics), while the acceptor was excited at 635-nm using a diode laser (WSTech, TECRL-25GC-635). Alternating-laser excitation (ALEX) of the sample was performed by synchronous modulation of the two laser sources to achieve alternating 50-µs periods of donor and acceptor fluorophore excitation, respectively. Laser intensities of 10 kW/cm^2 and 3.6 kW/cm^2 at the sample were used for exciting the donor and the acceptor fluorophores, respectively. FAD measurements were performed on the same microscope, by exciting Atto488 at an average intensity of ~0.14 kW/cm^2 at the sample. FCS measurements were performed on a separate custom-built fluorescence microscope described elsewhere (55), where Atto488 was excited using a 488-nm diode laser (TECBL-488nm, WorldStarTech) with a power of ~5 µW at the sample.

FAD analysis. FAD monitors the rotation dynamics of the emission dipole of the dye. The "wobble-in-a-cone" model (30) was used to fit the experimental FAD data:

$$r(t) = r_0\left[(1-a)e^{-t/\rho_{dye}} + a\right]e^{-t/\rho_{seg}} + r_{inf} \quad (1)$$

where ρ_{dye}, ρ_{seg} are rotational correlation lifetimes of the dye and the protein segment, respectively, and a is the fraction of the slower (segmental) component. The uncertainties of the fitted parameters were estimated by taking the standard deviation of the fitting results of 3-7 data sets collected consecutively; if any of the fitting derived errors were larger, then this value was reported. The baseline offset r_{inf} accounts for the slow, global motion of the whole protein; this is typically very small for IDPs due to their high backbone flexibility (29), but it can increase upon binding to their targets. The half-cone angle θ describes the wobbling cone of the dye emission dipole (56):

$$a = \left[\frac{1}{2}(1+cos\theta)cos\theta\right]^2 \quad (2)$$

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

FCS analysis. In FCS, the fluorescence autocorrelation function for free Brownian diffusion of a single molecular species with multiple relaxation components is given by (57) :

$$G(\tau) = \frac{1}{N_{eff}}\left(1+\frac{\tau}{\tau_d}\right)^{-1}\left(1+\frac{\tau}{s^2\tau_d}\right)^{-0.5}\sum_i\left(1+a_i e^{-\frac{\tau}{t_i}}\right) \quad (3)$$

In equation (3), N_{eff} is the average number of molecules in the detection volume, s is the ratio between the axial and the lateral radii of the detection volume ($s = z_0/w_0$), and τ_d is the average diffusion time, which is related to the diffusion coefficient ($w_0^2 = 4D\tau_d$) and to the hydrodynamic radius R_H of the molecule via the Stokes-Einstein equation(58); a_i and t_i are the amplitude and the lifetime of the ith relaxation component, respectively. Fitting was optimized by varying the number of relaxation components, the goodness of fit (χ^2) and Akaike information criterion (AIC) were calculated for each fit. The addition of relaxation component was accepted if: the χ^2 parameter decreased, the AIC decreased, and the fitting residuals were featureless. The uncertainties of the fitted parameters were estimated using the Jacobian from Levenberg-Marquardt least-squares fitting. Prior to each set of measurements, a calibration dye (rhodamine 110) was used to estimate the s and w_0 parameters, typically ~8 and ~250 nm, respectively.

smFRET analysis. A custom MATLAB script based on the 'MLT' algorithm was used to identify fluorescence bursts and sort them into donor-only, acceptor-only and dual-labelled (FRET) populations (32). The FRET efficiency for each burst was calculated based on the number of detected photons in the donor (I_D) and acceptor (I_A) channels:

$$E = \frac{I_A}{I_A+\gamma I_D} \quad (4)$$

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

where γ (**Eq. 5**) is a correction factor for the difference in detection efficiencies of donor and acceptor channels (η_D and η_A) and quantum yields of the dyes(46) (Φ_D and Φ_A), and was determined as described previously(32):

$$\gamma = \frac{\eta_A \Phi_A}{\eta_D \Phi_D} \tag{5}$$

In addition, corrections were applied on both I_D and I_A to subtract background, spectral cross talk, and direct excitation of the acceptor. The background was obtained from a measurement of the sample buffer while the corrections for cross talk and direct excitation were derived from donor-only and acceptor-only bursts. A smFRET histogram was constructed from all bursts detected for a given sample, each histogram was fit to a single Gaussian distribution.

To estimate the FRET efficiency if 4E-BP2 was a featureless statistical coil, 5000 conformers were generated using Trajectory Directed Ensemble Sampling (TraDES)(59) with 100% coil sampling in accordance with the sequences in **Table S1**. The FRET efficiency for each conformer was calculated using the python library *LabelLib*(60) (see SI section **2.3**).

For each FRET construct, a family of τ_{DA}/τ *vs* E curves were generated based on a WLC model with stiffness parameters κ between 0.01 and 1 using the *FRETlines* Python library(52). The chain stiffness for given dataset was selected as the κ value of the τ_{DA}/τ *vs* E curve passing through the centroid of a 2D Gaussian fit of the experimental τ_{DA}/τ *vs* E histogram. The confidence interval was established from the range of WLC models that agree with the centroid from fitting within the experimental accuracy of determining the FRET efficiency, i.e., 0.02 (32).

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

AUTHOR INFORMATION

Corresponding Authors

Email: claudiu.gradinaru@utoronto.ca

Author Contributions

C.C.G. was responsible for overall project management and supervision. S.S. and Z.Z. expressed, purified and labelled proteins, performed measurements and analyzed data. A.B. and J.D. designed single- and double-cysteine 4E-BP2 mutants. T.T. performed smFRET simulations. S.S., Z.Z., and C.C.G. wrote the manuscript. J.D.F.-K. contributed to interpretation of results and critical revision of manuscript.

Competing Financial Interest

The authors report no competing financial interest.

ACKNOWLEDGEMENTS

This work has been supported by the Natural Sciences and Engineering Research Council of Canada (NSERC RGPIN 2017 – 06030 to C.C.G.), the Canadian Institutes of Health Research (CIHR FND-148375 to J.D.F.-K.) and National Institutes of Health (NIH 5R01GM127627-03 to J.D.F.-K.).

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

REFERENCES

1. P. E. Wright, H. J. Dyson, Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *Journal of Molecular Biology* **293**, 321-331 (1999).
2. V. N. Uversky, Intrinsically disordered proteins from A to Z. *The International Journal of Biochemistry & Cell Biology* **43**, 1090-1103 (2011).
3. P. E. Wright, H. J. Dyson, Intrinsically disordered proteins in cellular signalling and regulation. *Nature Reviews Molecular Cell Biology* **16**, 18-29 (2015).
4. C. Haynes *et al.*, Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. *PLoS Comput Biol* **2**, e100 (2006).
5. A. K. Dunker, I. Silman, V. N. Uversky, J. L. Sussman, Function and structure of inherently disordered proteins. *Current Opinion in Structural Biology* **18**, 756-764 (2008).
6. Julie D. Forman-Kay, T. Mittag, From Sequence and Forces to Structure, Function, and Evolution of Intrinsically Disordered Proteins. *Structure* **21**, 1492-1499 (2013).
7. V. N. Uversky, Natively unfolded proteins: A point where biology waits for physics. *Protein Science* **11**, 739-756 (2002).
8. J. A. Marsh, J. D. Forman-Kay, Sequence determinants of compaction in intrinsically disordered proteins. *Biophysical journal* **98**, 2383-2390 (2010).
9. A. H. Mao, S. L. Crick, A. Vitalis, C. L. Chicoine, R. V. Pappu, Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proceedings of the National Academy of Sciences* **107**, 8183-8188 (2010).
10. N. Sonenberg, A. G. Hinnebusch, Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* **136**, 731-745 (2009).
11. S. Tait *et al.*, Local control of a disorder–order transition in 4E-BP1 underpins regulation of translation via eIF4E. *Proceedings of the National Academy of Sciences* **107**, 17627-17632 (2010).
12. A. Bah *et al.*, Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature* **519**, 106-109 (2015).
13. A. Lazaris-Karatzas, N. Sonenberg, The mRNA 5'cap-binding protein, eIF-4E, cooperates with v-myc or E1A in the transformation of primary rodent fibroblasts. *Molecular and cellular biology* **12**, 1234-1238 (1992).
14. A. Lazaris-Karatzas *et al.*, Ras mediates translation initiation factor 4E-induced malignant transformation. *Genes & development* **6**, 1631-1642 (1992).
15. A.-C. Gingras *et al.*, Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism. *Genes & development* **13**, 1422-1437 (1999).
16. R. J. Dowling *et al.*, mTORC1-mediated cell proliferation, but not cell growth, controlled by the 4E-BPs. *Science* **328**, 1172-1176 (2010).
17. R. Colina *et al.*, Translational control of the innate immune response through IRF-7. *Nature* **452**, 323-328 (2008).
18. C. G. Gkogkas *et al.*, Autism-related deficits via dysregulated eIF4E-dependent translational control. *Nature* **493**, 371-377 (2013).
19. E. Klann, J. D. Sweatt, Altered protein synthesis is a trigger for long-term memory formation. *Neurobiology of learning and memory* **89**, 247-259 (2008).
20. S. Lukhele, A. Bah, H. Lin, N. Sonenberg, J. D. Forman-Kay, Interaction of the eukaryotic initiation factor 4E with 4E-BP2 at a dynamic bipartite interface. *Structure* **21**, 2186-2196 (2013).
21. D. Peter *et al.*, Molecular Architecture of 4E-BP Translational Inhibitors Bound to eIF4E. *Molecular Cell* **57**, 1074-1087 (2015).
22. G. J. Brunn *et al.*, Phosphorylation of the translational repressor PHAS-I by the mammalian target of rapamycin. *Science* **277**, 99-101 (1997).
23. J. E. Dawson *et al.*, Non-cooperative 4E-BP2 folding with exchange between eIF4E-binding and binding-incompatible states tunes cap-dependent translation inhibition. *Nat Commun* **11**, 3146 (2020).

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

24. G.-N. Gomes, C. C. Gradinaru, Insights into the conformations and dynamics of intrinsically disordered proteins using single-molecule fluorescence. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1865**, 1696-1706 (2017).
25. L. A. Metskas, E. Rhoades, Single-molecule FRET of intrinsically disordered proteins. *Annual review of physical chemistry* **71**, 391-414 (2020).
26. V. N. Uversky, "Biophysical Methods to Investigate Intrinsically Disordered Proteins: Avoiding an “Elephant and Blind Men” Situation" in Intrinsically Disordered Proteins Studied by NMR Spectroscopy. (Springer, 2015), pp. 215-260.
27. G. F. Schröder, U. Alexiev, H. Grubmüller, Simulation of fluorescence anisotropy experiments: probing protein dynamics. *Biophysical journal* **89**, 3757-3770 (2005).
28. S. Milles, E. A. Lemke, Mapping Multivalency and Differential Affinities within Large Intrinsically Disordered Protein Complexes with Segmental Motion Analysis. *Angewandte Chemie International Edition* **53**, 7364-7367 (2014).
29. N. Jain, M. Bhattacharya, S. Mukhopadhyay, Chain collapse of an amyloidogenic intrinsically disordered protein. *Biophysical journal* **101**, 1720-1729 (2011).
30. J. R. Lakowicz, *Principles of fluorescence spectroscopy* (Springer, New York, N.Y., ed. 3rd, 2006).
31. Z. Zhang, D. Yomo, C. Gradinaru, Choosing the right fluorophore for single-molecule fluorescence studies in a lipid environment. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1859**, 1242-1253 (2017).
32. G.-N. W. Gomes *et al.*, Conformational Ensembles of an Intrinsically Disordered Protein Consistent with NMR, SAXS, and Single-Molecule FRET. *Journal of the American Chemical Society* **142**, 15697-15710 (2020).
33. F. Huang, W. M. Nau, A conformational flexibility scale for amino acids in peptides. *Angewandte Chemie* **115**, 2371-2374 (2003).
34. Keum *et al.*, A conserved motif within the flexible C-terminus of the translational regulator 4E-BP is required for tight binding to the mRNA cap-binding protein eIF4E. *Biochemical Journal* **441**, 237-245 (2012).
35. A. P. De Silva, T. S. Moody, G. D. Wright, Fluorescent PET (Photoinduced Electron Transfer) sensors as potent analytical tools. *Analyst* **134**, 2385-2393 (2009).
36. S. Doose, H. Neuweiler, M. Sauer, Fluorescence quenching by photoinduced electron transfer: a reporter for conformational dynamics of macromolecules. *ChemPhysChem* **10**, 1389-1398 (2009).
37. H. Chen, E. Rhoades, J. S. Butler, S. N. Loh, W. W. Webb, Dynamics of equilibrium structural fluctuations of apomyoglobin measured by fluorescence correlation spectroscopy. *Proceedings of the National Academy of Sciences* **104**, 10459-10464 (2007).
38. S. Mukhopadhyay, R. Krishnan, E. A. Lemke, S. Lindquist, A. A. Deniz, A natively unfolded yeast prion monomer adopts an ensemble of collapsed and rapidly fluctuating structures. *Proceedings of the National Academy of Sciences* **104**, 2649-2654 (2007).
39. H. Blom, A. Chmyrov, K. Hassler, L. M. Davis, J. Widengren, Triplet-State Investigations of Fluorescent Dyes at Dielectric Interfaces Using Total Internal Reflection Fluorescence Correlation Spectroscopy. *The Journal of Physical Chemistry A* **113**, 5554-5566 (2009).
40. L. J. Lapidus, W. A. Eaton, J. Hofrichter, Measuring the rate of intramolecular contact formation in polypeptides. *Proceedings of the National Academy of Sciences* **97**, 7220-7225 (2000).
41. F. Krieger, B. Fierz, O. Bieri, M. Drewello, T. Kiefhaber, Dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding. *Journal of molecular biology* **332**, 265-274 (2003).
42. J. K. Lum, H. Neuweiler, A. R. Fersht, Long-range modulation of chain motions within the intrinsically disordered transactivation domain of tumor suppressor p53. *Journal of the American Chemical Society* **134**, 1617-1622 (2012).
43. R. R. Goluguri, S. Sen, J. Udgaonkar, Microsecond sub-domain motions and the folding and misfolding of the mouse prion protein. *eLife* **8**, e44766 (2019).
44. S. Sen, H. Kumar, J. B. Udgaonkar, Microsecond Dynamics During the Binding-induced Folding of an Intrinsically Disordered Protein. *Journal of Molecular Biology* **433**, 167254 (2021).

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.14.476386; this version posted January 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

45. S. A. Waldauer, O. Bakajin, L. J. Lapidus, Extremely slow intramolecular diffusion in unfolded protein L. *Proceedings of the National Academy of Sciences* **107**, 13713-13717 (2010).
46. E. Sisamakis, A. Valeri, S. Kalinin, P. J. Rothwell, C. A. Seidel, Accurate single-molecule FRET studies using multiparameter fluorescence detection. *Methods in enzymology* **475**, 455-514 (2010).
47. S. Kalinin *et al.*, A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nature Methods* **9**, 1218-1225 (2012).
48. B. Schuler, W. A. Eaton, Protein folding studied by single-molecule FRET. *Current opinion in structural biology* **18**, 16-26 (2008).
49. A. Borgia, P. M. Williams, J. Clarke, Single-molecule studies of protein folding. *Annu. Rev. Biochem.* **77**, 101-125 (2008).
50. I. V. Gopich, A. Szabo, Theory of the energy transfer efficiency and fluorescence lifetime distribution in single-molecule FRET. *Proceedings of the National Academy of Sciences* **109**, 7747-7752 (2012).
51. B. Schuler, Perspective: Chain dynamics of unfolded and intrinsically disordered proteins from nanosecond fluorescence correlation spectroscopy combined with single-molecule FRET. *The Journal of Chemical Physics* **149**, 010901 (2018).
52. A. Barth *et al.*, Unraveling multi-state molecular dynamics in single-molecule FRET experiments-Part I: Theory of FRET-Lines. arxiv:2107.14770 (2021).
53. H. Hofmann *et al.*, Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. *Proceedings of the National Academy of Sciences* **109**, 16155-16160 (2012).
54. C. Igreja, D. Peter, C. Weiler, E. Izaurralde, 4E-BPs require non-canonical 4E-binding motifs and a lateral surface of eIF4E to repress translation. *Nature communications* **5**, 1-14 (2014).
55. Y. Li, R. V. Shivnaraine, F. Huang, J. W. Wells, C. C. Gradinaru, Ligand-Induced Coupling between Oligomers of the M_2 Receptor and the G_{i1} Protein in Live Cells. *Biophysical Journal* **115**, 881-895 (2018).
56. D. Badali, C. C. Gradinaru, The effect of Brownian motion of fluorescent probes on measuring nanoscale distances by Förster resonance energy transfer. *The Journal of Chemical Physics* **134**, 225102 (2011).
57. E. Haustein, P. Schwille, Fluorescence Correlation Spectroscopy: Novel Variations of an Established Technique. *Annual Review of Biophysics and Biomolecular Structure* **36**, 151-169 (2007).
58. A. Mazouchi, B. Liu, A. Bahram, C. C. Gradinaru, On the performance of bioanalytical fluorescence correlation spectroscopy measurements in a multiparameter photon-counting microscope. *Anal Chim Acta* **688**, 61-69 (2011).
59. H. J. Feldman, C. W. Hogue, Probabilistic sampling of protein conformations: new hope for brute force? *Proteins: Structure, Function, and Bioinformatics* **46**, 8-23 (2002).
60. M. Dimura *et al.*, Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems. *Current opinion in structural biology* **40**, 163-185 (2016).